DENOISE PRE-TRAINING FOR SEGMENTATION NEURAL NETWORKS

Martin Kolarik
Doctoral Degree Programme (2), FEEC BUT
E-mail: xkolar54@stud.feec.vutbr.cz

Supervised by: Radim Burget
E-mail: burgetrm@feec.vutbr.cz

Abstract: This paper proposes a method for pre-training segmentation neural networks on small datasets using unlabelled training data with added noise. The pre-training process helps the network with initial better weights settings for the training itself and also augments the training dataset when dealing with small labelled datasets especially in medical imaging. The experiment comparing results of pre-trained and not pre-trained networks on MRI brain segmentation task has shown that the denoise pre-training helps the network with faster training convergence without overfitting and achieving better results in all compared metrics even on very small datasets.

Keywords: deep learning, denoising, neural network, pre-training, segmentation

1 INTRODUCTION

The image segmentation is a process of partitioning an image into multiple segments. This is practically done as labelling each pixel of the image with a corresponding class that it belongs to. It is one of the key problems of computer vision. Image segmentation of medical data is an important problem. Accurate segmentation of medical image data helps doctors with automating routine work of image labelling and subsequently correct diagnosis determining. [1]

Neural networks in medical image segmentation are usually trained in a supervised manner. This means that in order to train the network we must create a large annotated dataset. The labelling also has to be done by somebody with necessary knowledge of human anatomy and therefore this process is costly and ineffective. Example of corresponding MRI scan slice and an labelled mask can be seen in Figure 1.

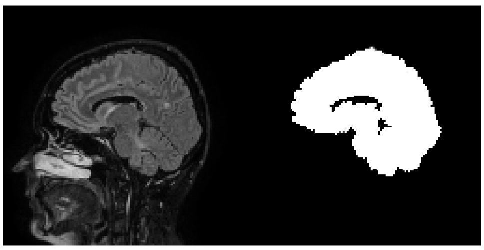

Figure 1: Example of sagittal MRI brain slice with a manually labelled brain segmentation mask from training dataset.

The proposed method is meant to help train deep neural networks using smaller datasets in an effective way. Using the denoise pre-training it is achieved better training of hidden layers in very deep architectures. Also the method helps the training algorithm to converge faster to a global optimum of the loss function.

2 RELATED WORKS

Pre-training of neural networks to achieve better results is a well known technique [2]. Using already pre-trained network is used to achieve better results when processing standard outdoor or indoor image scene, because available pre-trained networks are usually trained on such datasets. These pre-trained networks are therefore not suitable for medical image processing, because of the completely different character of the medical images and the data these networks were pre-trained on.

Segmentation of medical images can be done by classical mathematical methods [3] but convolutional neural networks usually achieve higher accuracy. First fully convolutional networks for image segmentation were used on indoor and outdoor segmentation because it was not easy to train them on small datasets [4]. Well explored and popular network architectures used for medical segmentation are based on autoencoder type network called U-Net [5]. The U-Net network is popular due to the ease of training the network from randomly initialized weights. Applying densely connected layers [6] to U-Net like architectures helps achieve much better performance [7] but also creates much larger networks. Training similar networks on very small datasets is not efficient because we are not able to properly train the deeper layers. Very good results have been achieved using denoise training in semi-supervised medical autoencoder image segmentation [8].

3 DATASET

Data used for this experiment consist of 9 MRI brain scans of patients age 35-55 years old scanned with flair sequence. Data have been anonymized and approved for scientific purposes. Data have been processed in resolution 128x128 pixels and each scan consisted of 256 slices. To simulate the experiment on small dataset I divided the data and performed the experiment with only 4 and then with 8 brain scans leaving the last 9th scan for testing.

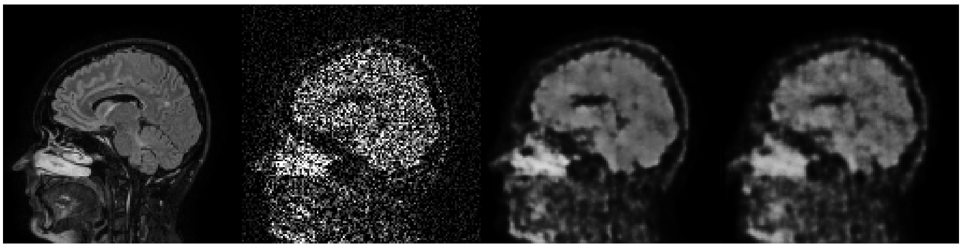

Figure 2: Example of brain slices used for training and denoised slices from the test scan. From left to righ - original training image, image with added 40 percent linear noise, image denoised using network trained on 8 scans, image denoised using network trained on 4 scans.

To prepare the pre-training data I added 40 percent linear noise to the training and testing data scans. Examples of original input image, image with added noise and denoised images can be seen in Figure 2. The pre-training process uses the noised scans as input images and tries to train the image to output the original image without noise in a supervised manner.

4 METHODOLOGY

The architecture of U-Net like autoencoder neural network used in the experiment can be seen in Figure 3. The experiment consisted of training the network on two datasets of different size - dataset of 4 scans and dataset of 8 scans to demonstrate problems with training such deep network on a small dataset.

I first trained the network to denoise the input image for 50 epochs on both dataset. Resulting weights were used as initialization weights for training the pre-trained networks. Both pre-trained and not

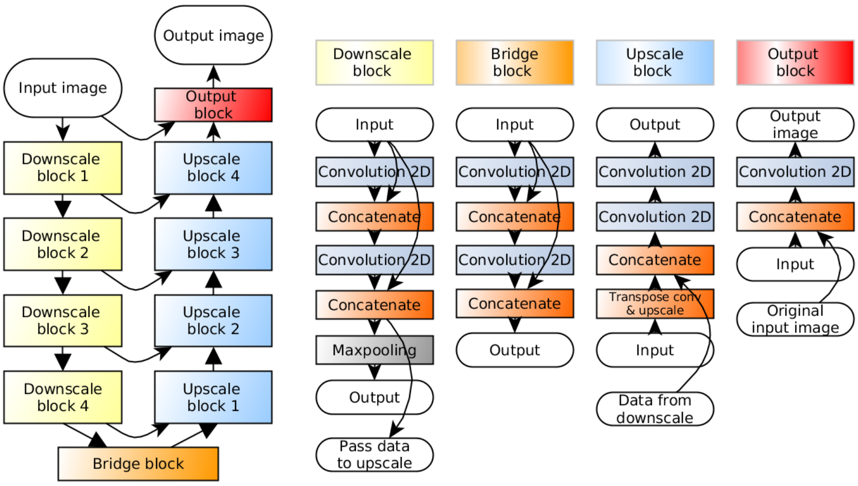

Figure 3: Architecture of the used Dense-U-Net neural network. [9]

pre-trained networks were then trained for 100 epochs each. Both networks started to show signs of overfitting after 100 epochs so there was no need for further training. The implementation was done in Keras [10] and I used Visceral segmentation tool [11] to calculate the resulting accuracy of the trained networks.

5 RESULTS

Denoised testing images from the pre-trained phase can be seen in Figure 2. Assuming the amount of noise, the network was able to learn the representation of the image data and the resulting weights were then used for training the pre-trained. Segmented brain models from the pre-trained and not pre-trained network trained on the larger dataset of 8 scans can be seen in Figure 4.

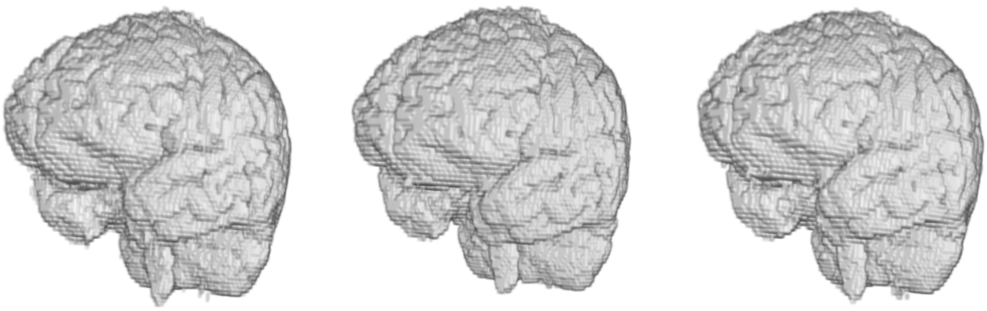

Figure 4: From left to right - 3D model of original labelled masks, 3D model from pre-trained network trained on 8 scans, 3D model from not pre-trained network trained on 8 scans. Notice that the pre-trained network achieved the least visible artifacts.

The pre-trained network was clearly able to better perform the segmentation without producing visible artifacts in the output masks. The development of the validation accuracy and the loss function is visualised in the Figures 5 and 6. There is clearly visible much faster convergence to global optimum of the pre-trained networks. There is also visible jump out of the optimum of the pre-trained 8 network during training. This was done by setting higher value of the parameter learning rate for pre-trained networks so they are able to find the global optimum of the loss function even when initialized on already trained weights. On the other hand this made the training process less stable.

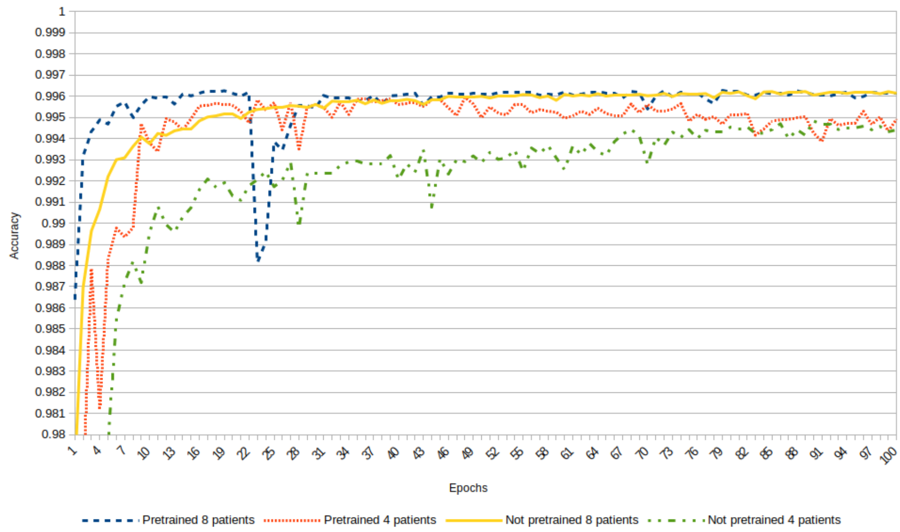

Figure 5: Visualization of the validation accuracy during training.

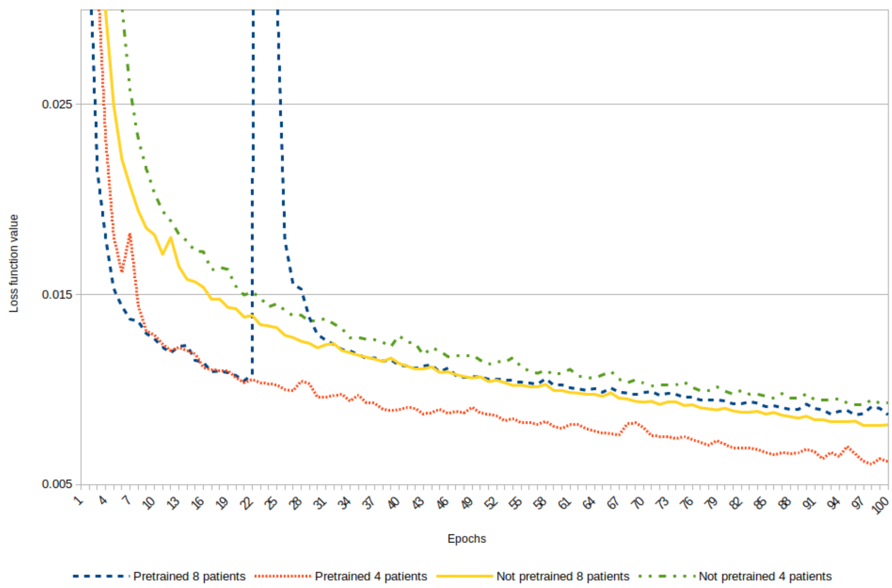

Figure 6: Visualization of the loss function during training.

In the Table 1 you can see results of comparison of accuracy, Dice coefficient and intersection over union of the all 4 trained networks. The results show clearly that the pre-trained networks were able to achieve better results in all metrics when compared to the network trained on the same size of dataset without pre-training.

Metric	**Pre-trained 8**	**Not pre-trained 8**	**Pre-trained 4**	**Not pre-trained 4**
P.A.	0.99626	0.99617	0.99609	0.99555
Dice c.	0.98239	0.98197	0.98153	0.97914
I.o.U.	0.96539	0.96458	0.96374	0.95914

Table 1: The accuracy, dice coefficient and Intersection over union comparison of the all 4 trained networks. All metrics were computed against manually labelled masks.

6 CONCLUSION

In this paper I have evaluated the effect of denoise pre-training for segmentation neural networks. This method has proven to be effective for faster training and achieving better accuracy of the trained networks. This method could be used as a standard procedure during training of neural networks for image segmentation not only for medical image processing especially for cases with small dataset.

REFERENCES

[1] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[2] D. Erhan, P.-A. Manzagol, Y. Bengio, S. Bengio, and P. Vincent, "The difficulty of training deep architectures and the effect of unsupervised pre-training," in *Artificial Intelligence and Statistics*, 2009, pp. 153–160.

[3] V. Uher, R. Burget, J. Masek, and M. K. Dutta, "3d brain tissue selection and segmentation from mri," in *2013 36th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, 2013, pp. 839–842.

[4] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[6] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[7] S. Jégou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 11–19.

[8] V. Alex, K. Vaidhya, S. Thirunavukkarasu, C. Kesavadas, and G. Krishnamurthi, "Semisupervised learning using denoising autoencoders for brain lesion detection and segmentation," *Journal of Medical Imaging*, vol. 4, no. 4, p. 041311, 2017.

[9] M. Kolařík, R. Burget, V. Uher, K. Říha, and M. K. Dutta, "Optimized high resolution 3d dense-u-net network for brain and spine segmentation," *Applied Sciences*, vol. 9, no. 3, p. 404, 2019.

[10] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[11] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, p. 29, August 2015.